DOCUMENT RESUME

ED 345 982 SO 022 073

AUTHOR Stahl, Robert J.; And Others
TITLE Achieving a High Test Score in the Social Studies Classroom: How Student Success Depends Heavily upon Unstated Criteria Used by the Teacher.
PUB DATE 14 Feb 92
NOTE 63p.; An extension of a paper presented at the Annual Meeting of the National Council for the Social Studies (Anaheim, CA, November, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; Cues; *Educational Research; Elementary Secondary Education; Higher Education; Preservice Teacher Education; *Social Studies; Teacher Influence; *Test Construction; Testing; *Testing Problems; Test Items; *Test Reliability
IDENTIFIERS Preservice Teachers; *Social Studies Teachers

ABSTRACT

Students in social studies classrooms frequently find themselves in situations where the information they need to provide in order to have a correct, complete response is not clear. For example, teachers frequently use cue words in questions and directions such as "identify," "describe," "analyze," and "explain." Such words are frequently used but rarely defined precisely in social studies classrooms. The use of such ambiguous terms forces students to guess what response is favored by the teacher. Thus, what students provide as answers on tests and what their scores are may be inaccurate representations of what they really know and have actually achieved. This report examines this hypothesis by focusing upon the test responses of pre-service teachers in social studies methods courses in order to illustrate how their own success on a test would be affected by the meanings and standards they and the grader used to evaluate their responses. It was found that among the pre-service social studies teachers, the terms used to cue responses to subject matter content test items were defined differently among the individuals involved and that these differences generated dissimilar expectations as to what is a complete, acceptable response to items where these cue terms are used. Implications of the findings for student success in social studies classrooms are drawn, including: if students' meanings for the identical term are different from those of their teacher and if they use their personal definitions to form a response to a test item, the chances are great that students will not be given full credit on those items regardless of how much time and effort they put into forming and writing the response. A list of 31 references is included. Four appendices containing information about the examination used in the report, responses given to the examination, and analysis of the data are included as well. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED345982

Achieving A High Test Score in the Social Studies Classroom: How Student Success Depends Heavily Upon Unstated Criteria Used by the Teacher

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ROBERT J. STAHL

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Robert J. Stahl
Associate Professor
Arizona State University
Division of Curriculum and Instruction
Tempe, AZ 85287-1911
602-965-7101
602-965-9144 (FAX)

George W. Chilcoat
Associate Professor
Elementary Education
Brigham Young University
Provo, UT 84602

Jerry Ligon
Associate Professor
Secondary Education and Foundations
Eastern Illinois University
Charleston, IL 61920

This report is an expansion of a paper entitled "Success on social studies tests is in the "mind of the grader:" Investigating differences in quality of answers to test items containing commonly used but rarely defined cuing terms" presented at the annual meeting of the National Council for the Social Studies, Anaheim, CA: November, 1990. As such the contents of this report supercede that paper of 1990.

(February 14, 1992)

SO 022 673

BEST COPY AVAILABLE

Achieving A High Test Score in the Social Studies Classroom: How Student Success Depends Heavily Upon Unstated Criteria Used by the Teacher[1]

Robert J. Stahl
George W. Chilcoat
Jerry Ligon

INTRODUCTION AND CONTEXT

Student success in social studies classrooms is dependent on much more than just paying attention, acquiring information, and effort. The quality of instruction, the pace of learning tasks, students' prior knowledge and entering abilities, extent and focus of student motivation, student perceptions of personal efficacy, and type and quality of learning activities are six factors that affect learning and contribute to each student's success (e.g., Bloom, 1976; Dempster, 1989; Carroll, 1989; Doyle, 1986; Slavin, 1989; Stahl, 1989, 1990a,b; 1982). Factors such as these influence what, how much, and how well students learn both within and outside social studies classrooms. While these factors influence what is learned on a day-by-day, course-by-course, and student-by-student basis, the extent of success for each student is measured and largely determined by his or her performance on tests and academic tasks.[2]

Tests: Definitions and Selected Roles

Ideally, within social studies classrooms, tests and academic tasks are directly aligned with targeted outcome knowledge and abilities described by curriculum and teacher objectives. Tests should serve to focus and channel student attention, effort, and learning. At the same time, tests should provide opportunities for teachers to gather information concerning the extent to which each student has attained or surpassed the knowledge and abilities expected (Airasian, 1991; Stahl, 1989). In addition, tests and their evaluation affect students' feelings of competence and influence how students approach personal study (Crooks, 1988). While tests perform these and many other

[1]The authors appreciate the editorial assistance of Micheal Jackson, Mountain View High School, Mesa, AZ in preparing this report.

[2]This of course assumes that teachers determine each student's final grade on actual performance rather than perceived notions about the student's success. It also assumes that the student's grades are not arbitrarily determined by "curving" grades and other artificial devices to inflate grades over actual performance measures. In this study, a test is "any device which is used or may be used to obtain information concerning the ability, condition, status, or quality of an entity, performance or phenomenon at a particular moment relative to one or more criteria or standards" (Stahl, 1983, 1989).

roles, the scores and grades students receive relative to their actual responses indicate how successful they have been in their learning efforts. Conventional wisdom assumes that the higher the score or grade, the more successful the student has been -- and is. Formal tests are clearly the major device secondary teachers use in order to measure and assess student success (e.g., Gullickson, 1985).

To obtain information relative to student learning, teachers ask questions, assign students tasks to complete, and provide directions and instructions for students to fulfill. Students are expected to use the information in the questions, assignments, and instructions as clues to guide their responses in providing evidence of their thinking and achievement. Students should have a clear notion of what they need to do or provide in order to do their best. They look to cues in the questions to clue them as to what is needed. However, what would the consequences be for student success were the cues provided not defined to provide a clear set of expectations for them in test situations?

Clarity and precision of expectations

Extensive planning should result in every test item clearly informing students as what they must provide to earn full credit for that item. Cuing terms should be clearly defined so that students and teachers alike have a consistent framework for how answers are to be constructed and written for full credit. For instance, if the teacher asks "In what year was President J. F. Kennedy assassinated?," students are expected to provide the year "1963." In this instance, the information in the question makes it quite clear what specific information students need to provide to earn full credit. Such precision eliminates any ambiguity as to what the teacher is expecting and what students need to provide to be successful in that situation. At the same time, this precision does not place students in the situation of having to guess just what information the teacher wants and will accept for them to be successful.

When given vague or ill-defined expectations, however, students must rely on various survival strategies to direct searches for useful cues in the questions and instructions provided. Ambiguous clues may be poorly deciphered or missed altogether. At other times, such cues may lead students to responses quite different from what the teacher expects and what the student, with more precise cues, could have provided.

Unfortunately, the precision illustrated in the above example is contrary to what many students typically find on worksheets, task assignments, tests and teacher generated instructions. For instance, were the teacher to asked, "When was President John F. Kennedy assassinated?", would the correct answer be "1963?"; "November, 1963?"; "November 22, 1963?"; or "on a Friday afternoon, November 22, 1963?" Suppose, however, a student answered "Before I was born," "On the day he died," "In the 1960s," or "After World War II ended." While these are all correct answers to

the question as it was actually asked (i.e., worded), few teachers would likely accept such statements as deserving full credit for an appropriate answer to this "when" question.[3],[4] The typical teacher fully expects students to "know" exactly what information will and won't be accepted in situations where the "when" question is asked. That sometimes the same teacher wants only a year, at other times a complete date, and at still other times information such as time of day, a particular month, day of the week, or even period of one's life, makes what is an acceptable answer to such questions very uncertain if not an absolute mystery for most students.

Students frequently find themselves in situations where what information they need to provide in order to have a correct, complete response is not clear. The cue words teachers use in questions and directions include terms such as "identify," "describe," "analyze," and "explain" -- words that are *frequently used but rarely defined precisely* in social studies classrooms. Unfortunately, these words are only a sample of the many ambiguous cue terms that continually find themselves being used in classroom situations. If a student is to succeed, he or she must often ignore the cues actually included in the question and rely upon guesstimating correctly what the teacher most likely wants as an appropriate response at that moment.

Students who find themselves in this "survival context" soon realize that their success is largely dependent upon how well they can invent and construct meanings for the cue terms being used; how well they can retrieve information and then construct responses that are consistent with their invented meanings for the cue terms; and how consistent their invented meanings are with those meanings their teachers will use in assessing the quality of their responses.[5] In such instances, their academic focus is not directed toward demonstrating social studies achievement, but on inventing ways to be successful on that test.

Academic success for students in the social studies presupposes that their teachers actually (a) have an explicatable and precise meaning for each term and for the type of information each term

---

[3]We have found that different teachers would give full credit for each of these different answers. For instance, while certain teachers gave full credit for "1963," others gave only half or no credit for "1963" and full credit only for "November 22, 1963" for the exact same "when" question.

[4]The use of this question that calls for the recall of a specific date is used to illustrate the need for clear cues regardless of the information being sought in the answer. This example in no way suggests that the authors endorse this type of question as being the most important or suitable one for social studies classrooms.

[5]Another, perhaps more vivid and relevant situation where this invention occurs may provide a clearer illustration for the in-service educator. A number of university students have told the researchers that during many of their university courses they depend heavily upon the feedback their instructors give after the first test in order to determine how they need to answer test questions on all future tests. They then hope their instructors continue to use this same criteria for responses and evaluating for all future tests. Comments by students such as "If I knew that's what he wanted for the answer to that question . . . ," and "The way her question was worded certainly didn't lead me to giving the kinds of information she accepted . . . ," reveal that the ambiguity mentioned here about precise cues and clear agreed-upon meanings for critical terms used as cues in questions and instructions is not limited to pre-college level classrooms.

requires in the responses students must provide, (b) used this meaning to frame the test item; (c) will use this meaning to determine an acceptable response for each particular item; and (d) will communicate and/or negotiate clear meanings for these terms so students can construct responses to meet the criteria the teacher will use. Teachers may provide these meanings on their own or work jointly with students to arrive at an acceptable set of definitions. The extent to which the teacher fails to do any of these four increases the probability that student success is determined arbitrarily by their teachers. In these instances, the degree of success in the form of a score or grade is dependent upon ambiguous, unshared, and often arbitrary criteria. Putting it more bluntly, *in many classrooms, what students provide as answers and what their scores are may be very inaccurate representations of what they have actually achieved and should receive as a score or grade.*

Social studies educators on all levels continue to use ill-defined and usually ambiguous terms to label expected student processes, outcomes, and responses. These educators remain unaware or have ignored the problems generated for students by unclear definitions for each of the abilities they require students to exhibit. Part of the reason for this failure may be that there has not been sufficient evidence of the kinds of affects ambiguous terminology may have on student achievement. In the interviews we conducted, the majority of teachers did not believe this lack of specificity generated any real problems for most students and that time spent developing, teaching, and using the definitions for these process and response terms would not result in any major changes in student responses.

This study was conducted, in part, to investigate the extent to which a student's success on a written test would be affected by the use of unstated definitions and standards to assess their individual performances. This report focuses on the responses of pre-service teachers in social studies methods in order to illustrate how their own success on a test would be affected by the meanings and standards they and a grader used to evaluate their responses. Hopefully, if teachers were to experience firsthand how the use of ill-defined terms and personally-invented meanings affected their own test scores, then they might comprehend what secondary students must deal with each time they respond to test items that include these same terms. This report represents the initial findings of an on-going study.

Purposes of this study

The purposes of this research were:

(a) to determine how pre-service social studies teachers respond to test items that include commonly used but rarely defined cue terms serving as guides or clues for framing appropriate test item responses;

(b) to determine the extent to which pre-service social studies teachers agree on what the appropriate responses are to the identical set of test items;

(c) to investigate how pre-service social studies teachers define selected cue terms that are commonly used in test items and task assignments relative to their measurement of student academic success in social studies classrooms on the pre-college level;

(d) to investigate pre-service teachers' reactions to information concerning differences in definitions they held and for commonly used but rarely defined cue terms, including the perceptions of the probable implications of these differences for their own teaching and for student learning, test responses, and test scores.

Purpose b was also the focus of an initial investigation in this same area with a small volunteer population of secondary school social studies students. A brief overview of the preliminary findings from the responses of these students is provided.

Finally, possible implications of the similarities and differences among teachers' definitions for these cue terms and of the use of these definitions as criteria for success for students enrolled in pre-college social studies courses were explored.

PRESERVICE SOCIAL STUDIES TEACHERS

Procedures and Methods

Pre-service teachers enrolled in five undergraduate social studies methods courses at Arizona State University, Brigham Young University, and Eastern Illinois University volunteered to participate in this study. The separate classes were made up of either all elementary level or secondary level pre-service teachers.

Students in these classes were provided a xerox copy of the first part of Chapter 21: "Settling the Last Frontier (1865-1890)" from Triumph of the American Nation (Todd & Curti, 1990: pages 486-496). They were also given a three page "Examination" as a required take home, open-book test. [See Appendix A for a copy of this exam.] The cue terms used in framing the 22 exam items were selected from among a list of cue terms one of the researchers had collected over a five year period from an assortment of teacher, commercial publisher, and author developed tests, worksheets, discussion questions, and directions. Specific cue terms included "identify," "describe," "analyze," "explain," and "compare." The cue terms "who" and "when" were also used in selected items.

The targeted cue terms were not defined for any participant prior to his or her completion of the take-home examination. Participants were instructed that this take-home exam was to be completed and that the exam would be graded. They were directed to answer all questions as correctly and completely as possible so that, when the exam was graded, they would earn full credit for their responses. These students were instructed to bring three copies of their completed assignment to class on the next day the class met.

BEST COPY AVAILABLE

During the next class period, each participant was given the anonymous exam papers of two of his or her peers in the same class. Once the papers were distributed, each participant was asked to read the two exam papers and to "grade it" as though these were turned in by students in an elementary or secondary class they might teach. Participants were directed to assign a number score of from "5" to "0" for each of the 22 items: a "5" was to be assigned to responses they considered to be correct and complete for each item while a "0" meant that the response was totally incorrect and unacceptable such that it did not merit any credits. Scores of between "0" and "5" were to be assigned for partially correct responses. No further details regarding points or standards of correctness were provided. Participants were also informed that, if they wanted to do so, they could use their copies of the text chapter to look up information relative to the test items.[6] Participants were then given time in class to score each item and to calculate the total points out of the 110 maximum each exam could earn.

The 11 students in one methods class were also directed to assign a letter grade to each exam they had just evaluated and scored. This task was assigned after they had all computed total test scores for the two exams. These students were directed to assign a letter grade of A, B ,C, D, or E and a "+" or "-" should one of these be appropriate. They were directed to assign the grade based upon the criteria they would use if the exam papers they received were from students in high school classes they might teach. For data analysis purposes, these letter grades were converted to numerical equivalents. A value of 4.0 was assigned for an "A; "3.7 for an "A-;" 3.3 for a "B+;" and so forth. This procedure would provide data as to how consistent these pre-service teachers would be in assigning similar grades to similar scores.

A total of 84 students in the five methods classes completed all of the tasks. Papers of students who brought in unreadable exams (due either to their own handwriting or to the poor quality of their reproduced copies) or who failed either to complete the scoring exercise or to record legible scores for all 22 test items were excluded from the data analysis part of the project.

These procedures enabled the researchers to examine the extent to which the responses of all participants varied among one another for each particular item and for the overall test. By having each paper evaluated by two individuals, the differences in the criteria for an acceptable response between two separate graders for each item for the same test could be studied.

The next section reports preliminary findings relative to the success these pre-service teachers attained on this take-home examination.

---

[6]Nearly all participants were observed to have referred back to the chapter at some time during the time set aside in the class period to complete the grading task.

Findings for Exam Scores of These Preservice Teachers

An initial set of 24 variables was generated. For each of the 22 test items, a score of 0-5 was recorded. A separate total score for items 1-20 (for a total of 100 points) and for all 22 items (i.e., 110 points for the total test score) were computed because of the possible responses that may have been generated in reaction to items 21 and 22.[7] (The reason for this possibility will be elaborated later.) Scores for each of these variables were available for each exam as graded by the two graders. For the 11 teachers in one class, a numerical equivalent to a letter grade was generated as a 25th variable.

Table 1 provides a sample of the actual responses to two of the 22 items by randomly selected participants. The numbers in the parentheses following each response in Table 1 represent the scores assigned to that particular response by each of the two peer "graders." For instance, if a (3)(5) occurs, one grader assessed this response as deserving 3 of 5 points while the second peer grader assessed it as deserving all 5 points. [See Appendix B for samples of more student responses to all 22 items.]

Table 1 about here

Table 2 lists the mean, standard deviation, and maximum scores for each of the 24 variables for these 84 pre-service teachers [See Appendix C for more data related to these scores.]. These data include the scores on both copies of each participant's exams. For test items 1 - 22, the maximum score was a "5." A maximum score of "100" was possible for the SUM20 variable (i.e., sum of items 1 through 20) and "110" was possible for SUM22 (i.e., sum of items 1 through 22). Given the nature of the assignment and that it was a take home exam that would be graded, the variation in scores for each item and for the total test score is quite remarkable.

Table 2 about here

The mean score for the 168 exams graded (i.e., two scores per 84 participants) was 94.2 (s.d. = 13.1) of a maximum of 110 or 85.7% correct among these preservice teachers on this exam.

---

[7]Test item 21 asked students to provide information that would indicate that they had an "understanding" of the information in the chapter. Item 22 asked them to reveal that they had "comprehended" this same information. Because of the likelihood that some students would claim that there was no difference between "comprehension" and "understanding," the researchers decided to compute two separate "total scores." One such score would be the total of the 20 items excluding items 21 and 22: the second score was the sum of the 22 items. The tables included in this paper include data relative to both of these totals. However, unless indicated otherwise, any reference to a "total test score" in the narrative refers just to the sum of the 22 items.

Table 1.

A sample of the responses of social studies methods students to and the scores of the two peer graders on two items on the 22-item take home examination.

Who was Chief Joseph?

a) He was the leader of the Nez Perce Indians. (5)(3)

b) Leader of the Nez Perce Indians. Led his people away from their homeland. Surrendered to the white man. (5)(5)

c) Leader of the Nez Perce Indians whose tribe, after trying to run to the Canadian border, was surrounded by American troops who wanted to relocate them on a reservation outside the fertile Great Plains. (4)(5)

d) Chief Joseph as the leader of the Nez Pierce Indians, a Pacific Northwest tribe. (3)(3)

e) Leader of a group of Indians known as the "Nez Perce." They were a peaceful tribe who had lived where the present day states of Oregon, Washington, and Idaho meet. Chief Joseph surrendered to the U.S. Army in 1877 in Montana near the Canadian border. His people were freezing to death. Chief Joseph had been attempting to lead his people to a free land in Canada. (5)(5)

f) Leader of the Nez Perce Indians who choose to run rather than fight the U.S. caverly. (5)(3)

g) Chief Joseph was the leader of the Nez Perce Indians who had lived on the Great Plains until pioneers invaded the land. He became famous for his attempt to flee to Canada and ultimately surrendered to U.S. troops very close to the Canadian border. (5)(4)

h) Chief Joseph was the leader of the Nez Perce Indians in the 1870s. He attempted to lead his people to a new homeland in Canada when the United States government ordered them out of their original lands. (5)(3)

Describe: Chief Sitting Bull.

a) The leader of the *Cheyenne* Indians he was described as honest, able, and idealistic. (1)(5)

b) Able, honest, and idealistic Sioux leader and chief. Led the Sioux over Custer in the Battle of the Little Big Horn. (5)(5)

c) Able, honest, idealistic. I assume he's Cheyenne, but not sure. (4)(3)

d) Chief Sitting Bull was the leader of the Cheyenne warriors. He was able, honest and idealistic. (5)(4)

e) He was the able and honest and idealistic leader of the Cheyenne warriors who fought Custer at the Battle of the Little Big Horn in Montana in 1876. (5)(3)

f) Able, honest, idealistic. (5)(4)

g) Has been described as an outstanding warrior considered to be honest and idealistic who aided in the defeat of Custer at Little Big Horn but who eventually was defeated by the U.S. Cavalry troops. (5)(5)

h) Sitting Bull was an able, honest, and idealistic leader of the Sioux Indians, one of the two leaders who played a major role in the defeat of General Custer at the Little Big Horn. (5)(5)

Table 2.

Mean, standard deviation, and maximum scores assigned by the two peer graders for the individual items and total test for the 84 preservice social studies teachers in the five methods classes.*

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.58	0.82	5.00	0.06
Q2	4.29	1.07	5.00	0.08
Q3	4.42	1.03	5.00	0.08
Q4	4.48	1.03	5.00	0.08
Q5	4.44	0.84	5.00	0.06
Q6	4.39	0.94	5.00	0.07
Q7	4.35	0.95	5.00	0.07
Q8	4.23	1.06	5.00	0.08
Q9	4.27	1.12	5.00	0.09
Q10	4.25	1.16	5.00	0.09
Q11	3.94	1.43	5.00	0.11
Q12	4.63	0.81	5.00	0.06
Q13	4.37	1.31	5.00	0.10
Q14	3.98	1.64	5.00	0.13
Q15	4.63	0.81	5.00	0.06
Q16	4.72	0.71	5.00	0.05
Q17	3.74	1.67	5.00	0.13
Q18	4.45	1.15	5.00	0.09
Q19	4.36	0.88	5.00	0.07
Q20	4.35	1.07	5.00	0.08
Q21	4.27	1.36	5.00	0.11
Q22	3.10	2.10	5.00	0.16
SUM20	86.86	11.50	100.00	0.89
SUM22	94.23	13.10	110.00	1.01

*The total number of scores for each variable was 168, representing two scores for each of the 84 participants.

A t-test, used to determine whether the scores for the identical items by the two graders were essentially the same, found a statistically significant difference ($p < .001$) in the scores assigned by the two peer graders for all 28 variables examined (see Table 3).

Table 3 about here

The mean GPA assigned to the scores by the 11 teachers in the one class was 3.27 (s.d. = .63), with the lowest grade assigned being a "C" attached to a score of 70 of 110 points (63.6%) and an "A+" assigned to a highest score of 106 (96.4%). Table 4 lists the scores assigned in this class by the peer graders and the letter grade given to those scores. The mean score for the 22 exams evaluated in this class was 90.9 (s.d. = 9.29) or 82.7% of the 110 points possible.

Table 4 about here

The extent to which the difference between the two overall test scores for the same participant's exam (i.e., sum of items 1-22) accurately reflect or actually misrepresent the real differences in scores across the 22 test items for the two graded copies of each student's exam was investigated To find out, the *absolute difference* between the two scores for each test item for the same exam was calculated. For instance, one grader may have assigned a score of "3" to the answer to the first test item and a "4" to the second. Meanwhile, the second peer grader may have assigned a "5" to the first test answer and a "1" to the second item on the identical exam. Note that there was an absolute difference of "2" points (3- 5 = -2) between the two scores for the first test item and "1" point (4 - 1 = 3) for the second item. All negative numbers (e.g., 3 - 5 = -2) were converted into positive numbers (e.g., the -2 becomes +2). This procedure allowed us to calculated the total amount of the differences for each test item and the 22 items combined. The maximum absolute difference possible for each test item was 5 points. Table 5 lists the mean, standard deviation, and range for the absolute difference between the two sets of scores for each of the 22-items and for the total test score for the 84 students.[8]

Table 5 about here

---

[8]The variable DSUM22 represents the difference between the total test scores for all 22 items by the two peer graders. The variable SMABDF22 is the difference between the total of the absolute differences between the two graders' scores for each of the 22 items for the same exam. DSUM20 and SMABDF20 are variables referring to sums for items 1 through 20 as mentioned previously. DGPA is the variable for the difference in the numerical equivalents to the letter grades assigned by the peer graders.

Table 3.

t-values for differences between separate scores assigned by two graders to identical answers for 22-individual test items and total test for the 84 preservice social studies teachers.

VARIABLE	t values	PR>|T|
Q1	5.97	0.0001
Q2	7.81	0.0001
Q3	6.83	0.0001
Q4	5.55	0.0001
Q5	7.16	0.0001
Q6	7.99	0.0001
Q7	7.81	0.0001
Q8	8.39	0.0001
Q9	7.05	0.0001
Q10	8.19	0.0001
Q11	7.66	0.0001
Q12	5.27	0.0001
Q13	5.80	0.0001
Q14	6.93	0.0001
Q15	5.55	0.0001
Q16	5.66	0.0001
Q17	7.33	0.0001
Q18	5.99	0.0001
Q19	6.24	0.0001
Q20	6.67	0.0001
Q21	5.09	0.0001
Q22	6.60	0.0001
SUM20	11.01	0.0001
SUM22	11.44	0.0001
SMABDF20	16.50	0.0001
SMABDF22	16.65	0.0001
DFABTO20	8.66	0.0001
DFABTO22	9.36	0.0001

Table 4.

The total test scores, percent of total test score, and letter grades assigned to the test scores by the 11 preservice teachers in one methods class.

TOTAL TEST SCORE	PERCENT	LETTER GRADE
106	96.4	A+
103	93.6	A
102	92.7	A-
100	90.9	A
99	90.0	A-
95	86.4	A-
95	86.4	B+
94	85.5	B
94	85.5	A
93	84.5	A-
93	84.5	B
92	83.6	B+
92	83.6	B+
91	82.7	B
90	81.8	B
90	81.8	B
85	77.3	B
83	75.5	B+
82	74.5	B
78	70.9	C
73	66.4	C+
70	63.6	C

Table 5.

Mean, standard deviation, and maximum absolute difference between two scores assigned to the same examination by two peer graders for the 84 preservice social studies teachers in the five methods classes.

VARIABLE	MEAN DIFFERENCE	STANDARD DEVIATION	MAXIMUM DIFFERENCE	STD ERROR OF MEAN
DQ1	0.61	0.93	3.00	0.10
DQ2	0.63	0.74	3.00	0.08
DQ3	0.55	0.73	3.00	0.08
DQ4	0.65	1.08	5.00	0.12
DQ5	0.55	0.70	2.00	0.08
DQ6	0.68	0.78	3.00	0.08
DQ7	0.64	0.75	3.00	0.08
DQ8	0.75	0.82	3.00	0.09
DQ9	0.88	1.14	5.00	0.12
DQ10	0.86	0.96	3.00	0.10
DQ11	0.95	1.14	5.00	0.12
DQ12	0.52	0.91	4.00	0.10
DQ13	0.52	0.83	4.00	0.09
DQ14	1.17	1.54	5.00	0.17
DQ15	0.54	0.88	4.00	0.10
DQ16	0.46	0.75	3.00	0.08
DQ17	0.88	1.10	5.00	0.12
DQ18	0.51	0.78	4.00	0.09
DQ19	0.61	0.89	5.00	0.10
DQ20	0.73	1.00	4.00	0.11
DQ21	0.48	0.86	5.00	0.09
DQ22	1.02	1.42	5.00	0.16
DSUM20	8.19	6.82	34.00	0.74
DSUM22*	8.67	6.94	34.00	0.76
SMABDF20	13.69	7.60	38.00	0.83
SMABDF22**	15.19	8.36	40.00	0.91

*DSUM22 represents the difference between the total test scores assigned by the two peer graders for the same exam.
**SMABDF22 represents the difference between the sum of the absolute differences between the scores on all 22-items as assigned by the two peer graders for the same exam.

The total test scores reported in Table 2 revealed a much closer agreement between peer graders than the actual absolute differences in scores they assigned for individual test items. The average difference between the two scores for the same exam was 8.67 points (s.d. = 6.94), with the largest difference being 34 points. As revealed in Table 4, when the absolute difference between scores for each test item were analyzed, the mean for absolute difference was 15.19 points (s.d. = 8.36). The largest absolute difference between the two sets of scores for the same examination was 40 points. The mean for the absolute difference nearly doubled that of the mean for the recorded final test scores (15.19 to 8.67) for the same exam. To illustrate this point, one student who had a 1-point difference between the total test scores assigned by the two peer graders had a 13-point absolute difference when the absolute difference points for the 22 individual items were sumed. Another student with a 7-point difference in the total test scores between the two graders was found to have a 29-point absolute difference across individual items. Even though there appeared to be a relatively high level of agreement among the peer graders for a identical test item responses when total scores were computed and then compared, in a majority of cases there were much larger differences between the scores assigned when absolute differences between each item rather than total exam score were contrasted

In order to describe the extent of the relationship between the differences between the total test scores by the two graders and sums of the absolute differences for the same 22 items, a Pearson correlation coefficient was generated. Across all 84 methods students, a coefficient of -.66 ($p < .0001$) was obtained. Coefficients of -.67, -.76, -.46, -.74, and -.83, respectively, were generated when scores for each of the five methods classes where examined separately. These negative coefficients indicate that the differences in total test scores by the two graders significantly misrepresent the extent of the actual differences that existed between the assessment of the 22 individual items by the two graders.

Tables 2-5 contain data that strongly suggest that, at least among these pre-service teachers, the two peers who graded each exam tended to hold and use quite different expectations for what was a highly acceptable response for each item and that their expectations were often very different from those of the individuals who wrote the exam responses. For these pre-service teachers, what was necessary for academic "success" for each item and for the overall exam was fully in the "eye and mind of each grader." This point is reinforced by the information received from each of the participants in one class who wrote prior to the grading task that he or she would give him/herself "110" of "110" points on the exam as he or she had answered it. The participants who earned "70," "73," and "78," respectively, on this exam were among those who claimed a perfect score prior to the peer grading exercise.

That these individuals would claim to have perfect papers is consistent with research literature describing how people tend to have confidence levels relative to the quality of their work that are frequently inconsistent with the actual quality of their ability, knowledge base, or performance (Glenberg, Sanocki, Epstein, & Morris, 1987; Moore, 1990; Weaver, 1990). These studies found little agreement between self-perceived levels of comprehension of specific sets of information and actual comprehension test scores. In other words, the "feelings of knowing" (Koriat, Lichenstein, & Fischoff, 1980) or "illusions of knowing" (Epstein, Glenberg, & Bradley, 1984) that individuals have about what they know or have comprehended are not very accurate or dependable indicators of what they actually "know" or "comprehend" as measured by objective performance tests. The findings in this study, although descriptive and not experimental, do lend support for the disparity cited above between perceived accomplishment and actual accomplishment as revealed by objective test data.

SECONDARY SCHOOL SOCIAL STUDIES STUDENTS

Procedures and Methods

The procedures applied during the pre-service social studies teachers study were identical to those followed by students in one secondary school class whose teacher volunteered to participate. The particular class was chosen because the class was going to study the same chapter in the same book used in the methods students study. The identical exam was distributed along with the same instructions for the take home assignment. Multiple copies of the completed exam were reproduced by the researchers for other steps of the study. The next day each student was given an anonymous copy of the exam of two of his or her peers. The exams which had only code numbers on them were randomly distributed. No student graded his or her own paper or knew the names of the students whose paper he or she graded. As with the teacher exams, these students scored each item from "0 - 5" and used their textbook to review any information that would help them decide what to accept and not accept for responses to each item. Of the 31 students assigned the set of tasks, 29 completed the tasks and turned in usable materials for the data analysis.

Table 6 provides the mean, standard deviation, and range of scores per each test item and for the overall test scores for both tests for these 29 students. These data also must be interpreted within the context that these students were given the take home assignment and told to complete it as correctly as possible since their responses would be graded.

Table 6 about here

Table 6.

Mean, standard deviation, and minimum and maximum scores by the two peer graders for the individual items and total test for the 29 secondary social studies students.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.71	0.70	2.00	5.00	0.09
Q2	4.29	1.12	0.00	5.00	0.15
Q3	4.29	1.24	0.00	5.00	0.16
Q4	4.45	1.08	0.00	5.00	0.14
Q5	4.34	0.91	2.00	5.00	0.12
Q6	4.47	0.78	2.00	5.00	0.10
Q7	4.38	0.99	0.00	5.00	0.13
Q8	4.03	1.21	0.00	5.00	0.16
Q9	4.34	1.02	0.00	5.00	0.13
Q10	4.12	1.39	0.00	5.00	0.18
Q11	4.21	1.22	0.00	5.00	0.16
Q12	4.64	0.87	0.00	5.00	0.11
Q13	4.40	1.34	0.00	5.00	0.18
Q14	4.14	1.46	0.00	5.00	0.19
Q15	4.67	0.69	3.00	5.00	0.09
Q16	4.67	0.73	2.00	5.00	0.10
Q17	3.78	1.68	0.00	5.00	0.22
Q18	4.55	0.96	1.00	5.00	0.13
Q19	4.52	0.80	2.00	5.00	0.11
Q20	4.29	1.08	1.00	5.00	0.14
Q21	4.55	0.94	0.00	5.00	0.12
Q22	2.83	2.23	0.00	5.00	0.29
SUM20	87.29	11.12	60.00	100.00	1.46
SUM22	94.67	12.27	65.00	110.00	1.61

When the differences between the scores given to the same exam by two peers were analyzed, many of the patterns found earlier in the methods students' data were found. The mean difference between the total scores for the same exam was found to be 10.45 points (s.d. = 8.47) with the smallest difference being one point and the largest being 34 points. When absolute differences in scores were examined, the mean absolute difference between the two scores for all 22-items was 15.00 points (s.d. = 9.19), with 39 points being the largest difference between scores assigned by the two peer graders. t-test analyses revealed the same pattern of significant difference for all variables examined as reported above in Table 4 for the preservice teacher data.

These findings reveal that among students in this high school classroom, exactly what was necessary for an appropriate, complete response to these test items varied widely from student to student. This difference could not be blamed on their having access to different information they were to study or different exam items to answer. The researchers concluded that a major contributing factor to these differences in test performance was that these high school students, like the preservice teachers, used their personally constructed notions of what the cue terms meant for them to do -- and they proceeded to develop their test responses and assess their peers' answers on the basis of these notions.

DEFINITIONS

Context and Procedures

How individuals define terms strongly influence how they interpret, assign meaning to, and make decisions in situations where these terms and their meanings are important (Stahl, 1983). In this study, it was assumed that when these participants were asked, for instance, "*to identify* the Battle of the Little Big Horn," their decisions regarding what information was appropriate or inappropriate were based on their conceptions of what it meant "to identify" something. In addition, their definitions would also tend to serve as the criterion for deciding the extent to which their own responses as well as those of their peers were correct, complete, and acceptable.

Those who completed the take-home exam and participated in the evaluation of the exams of two of their peers also defined 11 terms that often serve as cues in test items found in K-12 classrooms. In the order of their appearance on the definition assignment, these terms were: identify, describe, explain, understand, comprehend, compare, contrast, define, analyze, discuss, and cause. The term "discuss" was the only term not included as a cuing word in the items on the exam. This test and extended samples of the definitions these preservice teachers wrote for these cue terms are provided in Appendix D.

This exercise provided information about what these words meant to these pre-service teachers and high school students. Furthermore it enabled the researchers to collect information on the meanings these individuals had for these cuing terms and investigate the extent to which the definitions for individual terms were identical to or different from definitions of other terms on the same list. This procedure would also allow these participants to consider the differences among meanings they had for these terms and what these differences might might mean for testing and assessment of students they were to teach beginning as early as the next semester during their student teaching experience.

Findings Relative to The Meanings of The Cue Terms

The 84 sets of definitions were examined to determine the extent to which they are identical, similar to or different from one another. [See Table 7 for a sample of these defintions for three terms.]

Table 7 about here

Results of the initial investigation showed a remarkable degree of nonconformity in the definitions. For example, among the pre-service teachers, the term "identify" was defined as "locating information from a given source in order to provide reinforcement for learning," "to place or distinguish from others - to name," "to point out what is vital or important," "to write down important information about something but not as much as when you describe it," "to give a brief statement of who or what a person or event is," and "to locate and describe an idea presented." The pattern of widely diversified meanings was consistent throughout the 84 sets of definitions.

The definitions for the ten terms given by the 29 secondary school students differed as widely as those stated by the 84 pre-service teachers. This finding revealed that the meanings for these cuing terms were as varied and idiosyncratic among individual high school students as they were among these soon-to-be classroom teachers.

This phenomenon of differences among meanings is consistent with the psychological findings that learners at all ages and in every area invent and construct meanings to cues, terms, and everything else in their worlds (Stahl, 1983, in press; Wittrock, 1986). Unless individuals are assisted in acquiring a particular definition, each will invent and use his or her personally idiosyncratic meanings which will then viewed as being accurate. Such meanings may serve to be quite functional in typical situations. However, in situations where one person's meanings are to be used to assess the performance of another person, such as teacher-meanings used to assess student answers, it is important that these meanings be aligned. Therefore teachers and students must spend time ensuring their meanings match in those areas where such meanings are to be used to guide student

Table 7.

Sample of definitions provided by the 84 preservice social studies teachers to the cue words on the definitions test.

Identify
-- pinpoint person/place/thing and state what part they played in what is being studied
-- noting specific features
-- to pick out parts of a story or characters
-- to name or state who someone is or what something is
-- to catagorize
-- to categorize things
-- to know and explain
-- to show that you know what it is

Describe
-- list the attributes of
-- express something or something
-- explaining the setting and structure and plot of what happened as completely as possible
-- to give details which tells about someone, something or an event
-- explaining the details of items or ideas
-- give in detail
-- to explain in terms of appearances, characteristics of, etc.
-- to list and explain specific details about a specific item
-- show understanding by explaining what it is

Explain
-- using the attributes to show comprehension
-- discuss and write about a situation
-- discuss whatever happened as completely as possible
-- to give details, "reasons" to describe a concept
-- discussing the meaning or reason
-- discuss the meaning of something
-- to describe
-- to state in one's own words details about an event, object, or act.
-- discuss or write the meaning or reason for something

study and responses and to evaluate student achievement.

The area of most agreement of any of the cuing terms was for "cause." However, in reviewing the data concerning scores given for the item where participants were to list the cause of a particular event, both the pre-service teachers and high school students were found to differ significantly in the scores they assigned for the same answers by their peers. Even though in their own minds these participants had somewhat similar meanings for what constituted a "cause," they did not agree on what specific information given in the chapter constituted a correct and complete answer relative to the cause of the event cited. Indeed, the differences among their actual exam answers regarding the cause(s) of the event cited were found to be statistically significant ($p < .001$). Thus, having somewhat similar meanings for cue terms is not, in and of itself, sufficient to guide students to locate and provide equally similar and acceptable historical data congruent with the same cue terms. Furthermore, if specific causes cannot be determined given somewhat similar meanings and the identical historical data to study, then student success will rest almost entirely on the teacher's untaught criteria (i.e., meanings).

What These Differences In Meanings Meant to These Preservice Teachers

Not only did the pre-service teachers vary in their definitions of these 11 cuing terms, but there was also a wide disparity of opinion concerning the implications that this variability might have on their assessment of students being the next semester. Pre-service teachers in one secondary methods class were placed in small groups and asked to critiqued their test scores and explore reasons for the differences. After a lengthy discussion, nearly all agreed that the wide range of definitions and the failure to establish and then teach clear definitions to their students could only hinder their students' success. At the end of the group meeting, one wrote:

> "The implications of our differences [in test scores and definitions] are quite clear. Students would be confused, especially if I didn't use the same definition myself from test to test or question to question... It seems to me that it would be unfair to students to use these terms. If teachers in the field already know that these terms are unclear and thus are likely to have a negative affect on students, why haven't they decided to agree on definitions? They could at least do so in their own departments!!!!"

However, a number of preservice teachers advocated a very different point-of-view. Typical of their responses is the following: "For the majority of definitions, we were in general concurrence and in this regard we believe this [difference] would not provide for confusion on the student's part. The ones [definitions] that were different began very similarly, but in some cases went off in different directions, though the basis of the definitions were similar, i.e., some parts of the definition were the

same." Among the definitions these participants thought were in close agreement and would not lead to confusion in student thinking or responses were, for instance, for 'explain:' "to tell how or why something works, or happened the way it does/did," "to place into perspective a particular object, thought, or action," and "the ability to transfer information from one person to another." It was difficult for us to see these definitions as being sufficiently similar to one another so that an answer based upon any one would correspond well to answers generated by the other definitions.

Writing on behalf of a group, one preservice teacher concluded that "they [the group] did not believe any great problems [for students] would result from these minor definitional differences." [Underline emphasis and bracketed information added by the authors.] Quite remarkably, these teachers reached this conclusion after they had received less than perfect scores on their exams by two peer graders and were reminded that each of them had expected a perfect score on the exam as she or he had completed it.

These reports were especially interesting in light of the fact that nearly all of the pre-service teachers acknowledged in their groups that they themselves tended to assign scores for individual answers based upon what they put down for their own answers or on what they thought the test item asked for when they got set to grade the papers. Some reported they made their decisions about what was acceptable by comparing the answers on the two exams they were grading. In this latter case, they made a decision based not upon a single criterion but upon a comparison of the answers given on the two tests as they graded them. While they didn't think this was what a teacher should do in all cases, they considered it to be a "fair" way of evaluating answers when it wasn't clear what a complete and correct answer should entail. Yet, each of these preservice teachers strongly rejected the notion that this procedure was "fair" when their university professor suggested he might use these same non-defined cuing terms on exams and in evaluating these teachers' work in the methods course.

CONCLUSIONS, IMPLICATIONS, AND POSSIBLE COURSES OF ACTION

It is clear from these initial data that among these pre-service social studies teachers, the terms used to cue responses to subject matter content test items were defined differently among the individuals involved and that these differences generated dissimilar expectations as to what is a complete, acceptable response to items where these cue terms are used.[9] Furthermore, these data reveal that the differences among total test scores are likely to misrepresent the actual differences that exist in the quality of the responses to individual test items when considered collectively in the form of

[9]Preliminary data from one class of inservice social studies teachers revealed patterns of answers and grades very similar to those reported above for both preservice teachers and secondary school students.

absolute difference scores.

The implications of these findings to date for student success in social studies classrooms are many. The discrepancy among the definitions poses more than one problem for students. Four of these are touched upon here. First, if their meanings for the identical term are different from those of their teacher and if they use their personal definitions to form a response to a test item, the chances are great that students will not be given full credit on those items regardless of how much time and effort they put into framing and writing this response.

Secondly, if students are aware that their definition may not be the same as that of their teacher and if they desire to earn full credit for their responses to a particular item, then the students must spend time and effort during the test situation trying to invent meanings of cue terms that will allow them to generate the responses the teacher expects. If their invented definition is correct, students will likely be successful for that test item response at that moment. If their invented definition is incorrect, then their responses have little or no chance of being successful. In these instances, the focus of student energy is not on learning and revealing what they have learned or are learning. Rather, it is on surviving in situations where their success or failure as students depends on their ability to "read the mind" of their teachers for each test item on each test or assignment.

Thirdly, having somewhat similar meanings for terms may not be enough to guide individuals to locate or retrieve all of the particular information that the evaluator will expect as a totally acceptable response to a test item. In other words, being close in meaning is likely to produce different answers which in turn is very likely to result in different test scores.

Fourthly, these findings suggest teachers in a single department or school may find it valuable to develop clear and precise definitions to be used every time particular cue terms are included in questions or directions. Time should then be taken to help students comprehend and apply these definitions so that they become very skilled and consistent at using these meanings across situations. Students could then spend far more of their time and efforts in framing responses to meet clear expectations. Freed from the constant effort of "reading the mind" of the teacher as grader and evaluator, students can engage sooner and more fully in generating acceptable answers than they are able within ill-defined situations. Faculty members in Louisville, KY schools where such definitions have been developed and used report teachers and students alike have found these to be very effective tools for all concerned. Many are very positive about having taken the time to generate these definitions as expectations and then using them in every course taught. Most importantly, they are positive about how it has helped students academically by eliminating a lot of the confusion, failure, and apathy that often came from ambiguous meanings.

Even though much has been written about constructing test items and assessing students in social studies classrooms (e.g., Airasian, 1991; Brown, 1976; Moore, 1980; Williams & Moore, 1980), there is little evidence that these guidelines and notions are widely used by classroom teachers or

significant numbers of commercial publishers. While the items on the exam used in this study were cited as being "poor" by several critics at a recent national conference of the National Council for the Social Studies, the wording was lifted from samples of test items and activities written by and collected from a number of elementary and secondary classroom teachers, district test designers, and commercial materials within the past five years. The wording was typical of what we have found in analyzing many tests used on different levels of social studies instruction across a number of states. While all teachers do not write test items using such terms without providing clear definitions of what they require, these terms appear often enough in questions used for discussion, tests, and worksheets to warrant examination by teachers of the terms they use to cue students for learning success. Most importantly, we are not suggesting that these term be eliminated as cuing terms for student responses: we are advocating that for whatever terms are used, teachers and students have agreed-upon meanings that translate into clear expectations of what is needed in the response to receive full credit.

To date, a majority of educators at all levels and across all content areas have resisted the effort to develop a mutually agreed upon precise definitions for the cue terms they frequently use to guide, measure, and evaluate student learning. At the same time there is a major concern over what leads students to be "unsuccessful" and "unmotivated" in pre-college classrooms. As indicated in this study, the differences in definitions and meanings teachers and students have for the identical cue terms contribute significantly to the extent of student success or failure.

Considering the above, the chances are great that large numbers of students may not be doing as well as they could be in social studies classes because of the ambiguity of the meanings of the terms that are used to focus their learning, guide their responses, and assess the quality of those responses. While developing and using such definitions will not remedy every problem in these classrooms, this is one positive step every teacher can take to eliminate problems that result from this ambiguity. This notion is supported by the work of Bloom (1976), Guskey (1985), and others who report that students, including less able, slow, and even seemingly unmotivated ones, tend to be far more successful when they are clear as to precisely what they are to learn, how they are to learn it, and how they are to demonstrate what they have achieved.

Educators can take steps to examine the differences among the definitions that so greatly influence achievement and begin development of uniform parameters for each commonly used cuing term. In this way, test items can function to find out "what" students actually have learned relative to a clear set of expectations rather than revealing the results of students' efforts to "survive" in the arbitrary grading jungle of the social studies classroom.

REFERENCES

Airasian, P. W. (1991). Classroom assessment. New York: McGraw-Hill.

Block, B. H., Efthim, H. E., & Burns, R. B. (1989). Building effective mastery learning schools. White Plains, NY: Longman.

Bloom, B. S., (1976). Human characteristics and school learning. New York: McGraw-Hill.

Brown, L. B. (Ed.). (1976). Social Education (entire issue). 40(7, November-December).

Carroll, J. B. (1989). The Carroll Model: A 25-year retrospective and prospective view. Educational Researcher, 18(1), 26-31.

Commonwealth Center News (1990). Two studies of teachers' grading practices. 2(2, Spring), 1.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. Review of Educational Research, 58(4), 438-481.

Dempster, F. N. (1991). Synthesis of research on reviews and tests. Educational Leadership, 48(7, April), 71-76.

Dempster, F. N. (1989). The spacing effect: A case study in the failure to apply the results of psychological research. American Psychologist, 43, 627-634.

Doyle, W. (1986). Classroom organization and management. In M. C. Wittrock (ed.). Handbook of research on teaching. New York: MacMillan. pp 392-431.

Epstein, W., Glenberg, A. M., and Bradley, M. M. (1984). Coactivation and comprehension: Contributions of test variables to illusion of knowing. Memory & Cognition, 12, 355-360.

Glenberg , A. M. and Epstein, W. (1985). Calibration of comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 702-718.

Glenberg, A. M., Sanocki, T., Epstein, W., and Morris, C. (1987). Enhancing calibration of comprehension. Journal of Experimental Psychology: General, 116, 119-137.

Gullickson, A. R. (1985). Teacher perspectives of their instructional use of tests. Journal of Educational Research, 77, 244-248.

Guskey, T. R. (1985). Implementing mastery learning. Belmont, CA: Wadsworth.

Koriat, A., Lichtenstein, S., & Fischoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.

Moore, J. R. (1980). Social studies assessment: Current practice. In P. L. Williams and J. R. Moore (Eds.). Criterion-referenced testing for the social studies. Bulletin No. 64. Washington, D.C.: National Council for the Social Studies, 9-25.

Morris, C. C. (1990). Retrieval processes underlying confidence in comprehension judgments. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(2), 223-232.

Slavin, R. E. (1988). Student team learning: An overview and practical guide. Second edition. Washington, D.C.: National Educational Association.

Stahl, R. J. (in press) A context for "higher order knowledge:" An Information-Constructivist (IC) perspective with implications for curriculum and instruction. Part I. Journal of Structural Learning and Intelligent Systems.

Stahl, R. J. (April, 1990a). Learners as constructivists, infoschemata builders, and believers: A synthesis model of information processing with implications for thinking, learning, and instruction. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Stahl, R. J. (April, 1990b). What students need and do to become successful learners: An information-constructivist prospective of school learning." Paper presented at the annual meeting of the National Association for Research in Science Teaching. Atlanta.

Stahl, R. J. (April, 1989). Time alone does not mastery make: Extending Carroll's Model of School Learning in light of an information processing perspective." Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Stahl, R. J. (March, 1987). Ways of thinking about how humans think and learn: Practical models of information processing and learning. General Assembly paper presented at the Annual Conference of the Association for Supervision and Curriculum Development, New Orleans.

Stahl, R. J. (November, 1982). A synthesis of research on human information processing: Implications for education. Paper presented at the College and University Faculty Assembly during the annual conference of the National Council for the Social Studies, Boston.

Stahl, R. J., & Chilcoat, G. (November, 1989). Understanding in the social studies: A review of the meanings of understanding with implications for student learning and classroom instruction. Paper presented at the annual meeting of the College and University Faculty Assembly of the National Council for the Social Studies: St. Louis. (Submitted to ERIC)

Stahl, R. J., Chilcoat, G., & Ligon, J. (November, 1990). Academic success on social studies classroom tests is in the "Mind of the Grader:" Report of initial findings of an on-going study." Paper presented at annual meeting of the College and University Faculty Assembly of the National Council for the Social Studies. Anaheim, CA. (Revision submitted to ERIC)

Todd, L. P., & Curti, M. (1990). Triumph of the American nation. Orlando, FL: Harcourt Brace Jovanovich.

Weaver, III, C. A. (1990). Constraining factors in calibration of confidence. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(2), 214-222.

Williams, P. L., & Moore, J. R. (Eds.). (1980). Criterion-referenced testing for the social studies. Bulletin No. 64. Washington, D.C.: National Council for the Social Studies, (entire Bulletin).

Wittrock, M. (1986). Students' thought processes. In M. C. Wittrock (ed.). Handbook of research on teaching. Third edition. New York: MacMillan. pp 297-314.

APPENDIX A

THE 22-ITEM TAKE-HOME EXAMINATION

Examination

You are to use only the materials provided in the text pages you were given to answer the questions below:

1. Who was Chief Joseph?

2. What was the "Great American Desert"?

3. Describe the "Great Plains Indians."

4. What was the "new Indian policy"?

5. Identify General George Custer.

6. Describe the Battle of the Little Big Horn?

7. Identify: The Battle of Wounded Knee

8. Analyze the destruction of the buffalo.

9. Explain the advance of the railroads across the Great Plains.

10. What caused the "new Indian policy"?

11. Compare and contrast the Plains Indians and the white man in the Plains area from the 1830s until the end of the Indian wars.

BEST COPY AVAILABLE

12. Define: reservation

13. Identify: concentration

14. When did organized resistance against the white man end in the West?

15. Describe: Chief Sitting Bull.

16. Who was Chief Crazy Horse?

17. Analyze the movement of the settlers into the Plains region just before and soon after the Civil War.

18. Explain the events leading up to the Battle of Wounded Knee.

19. What the the most important theme or idea that was emphasized in this chapter?

20. Why is your answer to number 19 the most important theme or idea in this chapter?

21. It is expected that after you finished reading and studying this chapter that you understand what the chapter was all about. In the space below, write one or more paragraphs that will let me know that you <u>understand</u> the information in the chapter about this era in the history of the Great Plains region..

22. Use the space below to let me know that you <u>comprehend</u> the information in the chapter about this era in the history of the Great Plains region.

APPENDIX B

SAMPLES OF RESPONSES TO SELECTED ITEMS ON THE 22-ITEM EXAMINATION

Table B1.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Analyze the movement of the settlers into the Great Plains region just before and soon after the Civil War.

a) They first moved through the Plains on their way to the Pacific and described it as a great desert. But then began to realize they could farm and raise cattle there. (3)(5)

b) Prior to the Civil War (1848) settlers were very hesitant to move into or onto the Great Plains. Life was hard, lack of water, Indians, lawlessness, little game. As time went on settlers discovered how good and fertile the land was for farming and grazing. Yet the Indians were still prevalent. The government set up policies to allow more settlers to move in. Keeping the Indians concentrated. After the war, many soldiers from both sides needed work, so fighting Indians became an easy end to their means. Thus more and more moved in as the Indians became fewer and fewer. (4)(5)

c) Very sporadic, first on horseback, wagons, and then stagecoaches. After the Civil War, Great Plains attracted increasing numbers of land speculators, ranchers , miners, engineers, and farmers. These newcomers were determined to possess the land and its resources for themselves. (3)(4)

d) The movement of settlers into the Plains region was slow before the Civil War. They were threatened by harsh conditions, Indian attacks, and forms of slow transportation. As the stagecoach began to be used, more of the railroads built, the settlers were able to penetrate the territory more easily. (3)(3)

g) Prior to the Civil War most settlers believed the Plains to be uninhabitable. By the 1850s they realized their mistake as reports of good farm lands and cattle raising areas became apparent. The development of the railroads and the end of the Civil War were also factors in this expansion. (5)(4)

h) Before the Civil War, settlers were just beginning to move into the Plaions region, recieving reports that the region actually was good for farming and cattle raising. After the Civil War, increasing number of settlers moved into the area, not only farmers and ranchers, but also land speculators, miners, and engineers. (4)(5)

Table B2.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Describe the Battle of the Little Big Horn?

a) General Custer and 264 troops attempted to remove Cheyenne and Sioux Indians from that area. Sitting Bull and Crazy Horse were the Indian leaders and they killed Custer and *all his troops*. (4)(5)

b) General Custer and 264 troops attacked a large party of Sioux and Cheyenne Indians led by Crazy Horse and Sitting Bull. The battle occurred near the Little Big Horn River in Montana. All of Custer's men were killed. marked the end of major fighting in the northern Great Plains. (5)(5)

c) In June, 1976, Custer attacked a large camp of Sioux and Cheyenne near the Little Big Horn River in Montana. Sitting Bull and Crazy Horse defeated him and his 264 troops. Major humiliation. (4)(4)

d) The Battle took place in Montana. It consisted of the Sioux and Cheyenne headed by Sitting Bull and Crazy Horse who were fighting their removal to the reservation. They were successful in defeating the calvary, but ultimately settled on the reservation. (4)(5)

e) An attack lead in June of 1876 by General Custer against a large group of Sioux and Cheyenne Indians camped near the Little Big Horn River in Montana. The Indians were lead by two outstanding leaders, Sitting Bull and Crazy Horse. This became "Custer's Last Stand" as he and his who detachment of 264 troops were killed. (4)(5)

f) Custer attacked village. Indians led by Crazy Horse and Sitting Bull killed all Custer's men and mutilated bodies of calvary soldiers. (5)(5)

g) Battle of the Little Big Horn was a fierce battle in 1876 in which General Custer attacked Sioux and Cheyenne Indians along the Little Big Horn River in Montana and was defeated. (4)(4)

h) In this battle, General Custer and his troops attacked a large camp of Sioux and Cheyenne Indians in Montana and were all killed. The battle was a humiliation at the time for the governemtn, but brought a temporary end to fighting in the area as the Indian leaders from the battle were soon pursued and captured. (4)(5)

Table B3.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Compare and contrast the Plains Indians and the white man in the Plains area from the 1830s until the end of the Indian wars.

a) The Plains Indians roamed the plains hunting mainly buffalo. As they were forced onto reservations they resisted with great intensity. The white man on the other handsaw the plains as a place for white people to own land and live on. They felt the Indians must be concentrated on reservations. The Indians saw no ownership of land, no boundaries, the white man did. (3)(2)

b) The white man of the 1830s stayed to the east of the 100th meridian generally. The Plains Indians during this period were "veterans" of the land. They knew how to farm and hunt the "arid" plains. By the 1850s the notion that the plains were desolate wastelands was dispelled and more and more settlers kept moving westward. The Indians had an "edge" early in the century with their weapons , but as the revolver was invented and became more available, settlers were more compelled to move to the plains. As the Indians' resistance weakened, white men increased on the scene. The Indians were conserving, environmentally conscious people (buffalo), whereas the whites were wasteful greedy people. The Indians continued to lose land and the whites continued to gain land. Both farmed, grazed, and hunted on the land. (3)(5)

c) In the 1830s, the Plains Indians enjoyed relative prosperity and tranquility in the Plains area because the white man felt the land was worthless. However once that misconception was realized, the white man wanted their piece of the land, regardless of Indian considerations. This led to strife which lead to the Indian wars and eventually ended in the brutal massacre of Indians at Wounded Knee. (2)(4)

g) The white man and Indians seemed to trade superiority during this time. Both with the common goal of control of the land. In the 1830s the Indians were superior to the white man in many ways: adaptability to their environment, hunting, riding, as warriors. By the late 1830s and the invention of the revolver, white men steadily increased in power, breaking promises and forcing Indians onto reservations. (3)(4)

h) Early in this period, the Plains Indians recognized the abundance of the Plains area for subsistence, whereas the white man saw the area as an arid, uninhabitable desert. By the 1850s, the white man too recognized the value of the Plains for making a living, but still approached the task from a different perspective than the Indians. The Indians made their living off the existing resources of the Plains, the white man attepted to remove the main animal resource, the buffalo, and replace it with their own cattle, as well as introduce agriculture. (4)(5)

Table B4.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination

ITEM: When did organized resistance against the white man end in the West?

a) The Battle of Wounded Knee. (4)(5)
b) In 1886 organized resistance in the Southwest and Southern Plains came to an end with the surrenders of Geronimo in 1886. In the "west" in general, the Battle of Wounded Knee. (4)(4)
c) With the surrender of Geronimo in 1886. (5)(5)
e) After the Battle of Wounded Knee ended in South Dakota in 1890. (0)(1)
f) After the Battle of Wounded Knee (1889) although "battle" doesn't sound like the Indians were either "organized" or actively resisting. (5)(5)
g) Organized resistance against the white man came to an end in 1886 after three centuries of warfare when Geronimo surrendered. (5)(4)
h) Organized resistance against the white man ended in 1886, when the Apache Chief Geronimo surrendered. (4)(5)

Table B5.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Who was Chief Joseph?

a) He was the leader of the Nez Perce Indians. (5)(3)
b) Leader of the Nez Perce Indians. Led his people away from their homeland. Surrendered to the white man. (5)(5)
c) Leader of the Nez Perce Indians whose tribe, after trying to run to the Canadian border, was surrounded by American troops who wanted to relocate them on a reservation outside the fertile Great Plains. (4)(5)
d) Chief Joseph as the leader of the Nez Pierce Indians, a Pacific Northwest tribe. (3)(3)
e) Leader of a group of Indians known as the "Nez Perce." They were a peaceful tribe who had lived where the present day states of Oregon, Washington, and Idaho meet. Chief Joseph surrendered to the U.S. Army in 1877 in Montana near the Canadian border. His people were freezing to death. Chief Joseph had been attempting to lead his people to a free land in Canada. (5)(5)
f) Leader of the Nez Perce Indians who choose to run rather than fight the U.S. caverly. (5)(3)
g) Chief Joseph was the leader of the Nez Perce Indians who had lived on the Great Plains until pioneers invaded the land. He became famous for his attempt to flee to Canada and ultimately surrendered to U.S. troops very close to the Canadian border. (5)(4)
h) Chief Joseph was the leader of the Nez Perce Indians in the 1870s. He attempted to lead his people to a new homeland in Canada when the United States government ordered them out of their original lands. (5)(3)

Table B6.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: What was the "Great American Desert?"

a) The plains. (3)(1)
b) The area of the Great Plains. The white settlers were accustomed to trees and abundant water so when they got to the grassy plains it appeared like a desert to them. (4)(5)
(c) The Great Plains. Perceived by the Federal Govt as being arid and uninhabitable. However was a misconception and traders and "pioneers" wanted to settle there. (3)(3)
d) The "Great American Desert" was that area west of the 100th meridian ano east of the Rockies. It consisted of states such as the Dakotas, Nebraska, Wyoming, Kansas, and Oklahoma. (3)(4)
e) Also known as the "Great Plain," it reached east to west from the 100th meridian toward the Rocky mountains. A unifying factor of the area is the extremes of climate, including rain storms, tornadoes, hail, and wind. (4)(4)
f) Name given to the great plains in the early part of the 19th century. Name was based on mistaken belief that the great plains were arid and uninhabitable. (5)(5)
g) The Great American Desert was land designated by the 100th meridian (now known as the Great Plains) which was named during the early 1800s by explorers who considered the areas uninhabitable. (4)(4)
h) The Great American Desert was the name explorers and early settlers gave to the Great Plains in the first half of the 1800s. Since the region was not wooded it was given the title "desert," a dry, uninhabitable area. (4)(4)

Table B7.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Describe the "Great Plains Indians."

a) They hunted buffalo and were expert horse riders. They were also very good warriors. (3)(4)
b) They mainly consisted of previously "relocated" Eastern Indians that the government and whites in general "pushed" to an undesirable environment. These Indians were great hunters and farmers, and knew how to "break" wild horses and fight off intruders until the invention of the revolver. (3)(5)
c) Determined to defend their hunting grounds and their way of life. Were expert riders. Hunted buffalo which provided them with food, clothing, and shelter. Powerful adversaries. Rode expertly. Fought with spears first before the rifle. (5)(5)
d) The Great Plains Indians were groups of nomadic peoples that roamed the plains region on horses. They followed their food source, the buffalo, from which they made much of their supplies. (5)(4)
e) These were the tribes of Arapaho, Crow, Comanche, Mandan, Sioux, Arikara, Cheyenne, and Pawnee Indians who lived in the Great Plains area. They were highly skilled hunters who fought with arrows and long spears on horseback. But with the advent of revolvers, they were no longer a match for the white man. (4)(4)
f) Various tribes, often at war with each other, primarily nomadic followers of buffalo herds, characterized by their horsemanship and use of bow and arrow. (5)(5)
g) The Great Plains Indians can be described as expert riders, excellent hunters and and powerful adversaries who adapted themselves to the Plains and remained superior to white settlers until the 1830s. (5)(5)
h) The Great Plains Indians were tribes that had adapted to the plains environment for their living. They hunted on horseback the abundant plains buffalo as a source of food, clothing, and shelter. (4)(5)

Table B8.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: What was the "new Indian policy?"

a) It was a policy of "concentration," where the new Bureau of Indian Affairs confined Indian tribes to limited areas. (2)(4)
b) A policy developed by the BIA to concentrate the tribes into certain areas (i.e., reservations) so as to limit where the Indians could live and give whites traveling to the far west easier access. (4)(5)
c) 1849 - Bureau of Indian Affairs became part of the Department of the Interior and its policy was concentration in an attempt to confine the Indian tribes to certain limited areas of the West where they could live their own lives and hoped it would reduce warfare between tribes and clear them in areas for whites heading for California. (4)(5)
d) The new Indian policy was one of "concentration." The Department of Interior wanted to gather the Indians into a group so as to make more room for the white settlers. (4)(4)
e) In 1849 the Bureau of Indian Affairs became part of the Dept. of the Interior, and new policies toward the Indians emerged. One policy was "concentration" which hoped to assign tribes to a certain limited area so that the white settlers could have clear, peaceful routes to California and Oregon. (5)(5)
f) Was an attempt to confine Indians, by force of arms if necessary, to reservations and to thereby convert them from nomadic hunters to fixed plat farmers and/or dependents of U.S. Federal programs administrated by "Indian agents." (4)(5)
g) The "New Indian Policy" was an attempt by the Bureau of Indian Affairs to confine Indians to limited areas in the hopes of reducing tribal warfare and clear routes for pioneers heading west. (5)(4)
h) The "New Indian policy" was to greatly limit the range of Indian tribes on the plains and so opened much new land for settlement by whites and reduce the danger of Indian attacks on overland travelers. (4)(4)

Table B9.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Identify General George Custer.

a) Custer was an Army General who led troops against the Sioux and Cheyenne and was killed *with all his men* in the Battle of the Little Big Horn. (5)(5)
b) An experienced General and Indian fighter. He had a history of brutal massacres credited to him. Was the general who was defeated and killed at the battle of the Little Big Horn. (5)(5)
c) An experienced Indian fighter, leader of the 7th cavaly. Defeated at Little Big Horn. Several years earlier attacked a peaceful Cheyenne Indian village and killed unarmed women and children. (4)(5)
d) George Custer was the General of the 7th Cavalry. He is best known for his defeat by the Sioux at Little Big Horn. (5)(5)
e) A General in the Army who commanded the 7th Cavalry. He was an experienced Indian fighter. (3)(4)
f) Experienced Indian fighter who had massacred a peaceful Indian village on the Washita river. General slaughtered after attacking an Indian village on Big Horn River. (5)(5)
g) Custer was an "experienced Indian fighter" and leader of the failed U.S. Cavalry removal operation of the Sioux Indians out of the Black Hills. (5)(4)
h) General George Custer was an experienced, sometimes ruthless, Indian fighter who became famous for a battle he led his troops into against the Sioux and Cheyenne Indians. He and all of his troops were killed in the battle. (4)(5)

Table B10.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Describe: Chief Sitting Bull.

a) The leader of the *Cheyenne* Indians he was described as honest, able, and idealistic. (1)(5)
b) Able, honest, and idealistic Sioux leader and chief. Led the Sioux over Custer in the Battle of the Little Big Horn. (5)(5)
c) able, honest, idealistic. I assume hes Cheyenne, but not sure. (4)(3)
d) Chief Sitting Bull was the leader of the Cheyenne warriors. He was able, honest and idealistic. (5)(4)
e) He was the able and honest and idealistic leader of the Cheyenne warriors who fought Custer at the Battle of the Little Big Horn in Montana in 1876. (5)(3)
f) Able, honest, idealistic. (5)(4)
g) Has been described as an outstanding warrior considered to be honest and idealistic who aided in the defeat of Custer at Little Big Horn but who eventually was defeated by the U.S. Cavalry troops. (5)(5)
h) Sitting Bull was an able, honest, and idealistic leader of the Sioux Indians, one of the two leaders who played a major role in the defeat of General Custer at the Little Big Horn. (5)(5)

Table B11.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Who was Chief Crazy Horse?

a) The leader of the Sioux Indians, Crazy Horse was described as uncompromising, reckless and a military genius. (5)(4)
b) Chief and leader of the Sioux. An uncompromising, reckless military genius. Most honored hero of the Sioux. (5)(5)
c) Uncompromising, reckless, a military genius and most honored hero of the Sioux. (4)(5)
d) Crazy Horse was the leader of the Sioux tribe and he was uncompromising, reckless, a military genius, and the most honored hero of the Sioux. (5)(4)
f) uncompromising, reckless, honored hero among Sioux, and a military genius. (5)(4)
g) A Sioux warrior known as their most honored hero, considered to be a military genus but extremely reckless. He also figured prominently in Custer's last stand but then was defeated and forced by the U.S. to live on a reservation. (5)(5)
g) Chief Crazy Horse was a leader of the Sioux, their most honored hero, and one of the two leaders who played a major role in the defeat of General Custer at the Little Big Horn. (3)(5)

Table B12.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Identify: The Battle of Wounded Knee.

a) In 1870 a group of Sioux Indians were traveling in search of food. The Army intercepted them and disarmed them. A shot rang out and the troops opened fire on the unarmed Indians. 290 men, women, and children were killed. (4)(5)

b) Event that ended all organized armed resistance by Indians in the Great Plains. An "innocent" band of Sioux were halted or detained during a travel in which they were searching for food/protection. One shot caused an outbreak, killing many. Resistance caused by the Ghost Dance scare. (4)(5)

c) Should be the massacre at Wounded Knee. Where troops murdering 40 men and 200 women and children at Wounded Knee, South Dakota. Stopped organized resistance against Indians - 1890. (4)(4)

d) The Battle of Wounded Knee was a confrontation between the 7th Calvery and the Sioux who were on their way to Pine Ridge Reservation. Because resent fears about Indian resurgence they were surrounded by soldiers and killed when a disturbance broke out. (4)(3)

e) White settlers fearful of the "Ghost Dance" movement among the Indians, which called for an Indian messiah to appear, demanded that the Army put an end to this Indian activity. At Wounded Knee in South Dakota in December 1890, a unit of the 7th Cavalry arrested a band of Sioux Indians bound for the Pine Ridge Reservation. A disturbance broke out and the troops using gatling guns killed all the disarmed Sioux (90 men and 200 women and children.) This action ended all organized resistance from the Indians. (5)(5)

f) "Battle" in 1889 where 7th cavalry opened fire on group of unarmed Sioux men, women and children. (4)(4)

g) Battle of Wounded Knee was a massacre by U.S. Cavalry in South Dakota in 1890 of Sioux Indians which killed almost 300 -- mostly women and children. (4)(4)

h) The battle of Wounded Knee was a massacre of 290 unarmed Sioux Indians by government troops. This group of Indians was traveling peacefully to a reservation, but was apprehended by the troops in response to the fears of some settlers that an Indian uprising was imminent. (4)(5)

Table B13.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Analyze the destruction of the buffalo.

a) The expansion of the railroads divided the buffalo herds, but the main reason for the destruction was a policy of the government to reduce buffalo herds to keep the Indians under control. This policy encouraged the slaughter of the buffalo. (4)(5)

b) According to this text's version of the destruction, the buffalo were killed in order to "doom" the Indians, knowing that they were utterly dependent on them. For various reasons, the government condoned such actions as a means to "choke" the Indians. (3)(5)

c) Plains Indians depended upon them with blessing from Federal Govmt who wanted to confine them to reservations. Hunting parties killed them and used their hides for eastern markets. (3)(5)

d) The destruction of the buffalo took place in the 1870s. The purposeful anuilihilation of the herds was designed to keep the Indians on its reservations by depleting them of their source of goods and food. (5)(4)

f) As rail lines began to cut across the plains, migratory routes of buffalo were destroyed. Additionally, railroad brought whites who wantonly destroyed herd of buffalo, in part because they believed herd destruction would keep Indians on reservations. (5)(3)

g) The destruction of the buffalo was a deliberate intervention and encouragement by the U.S. Govt of hunters to destroy buffalo and eventually forced Indians out of control of the Plains. (5)(5)

h) The buffalo were destroyed by white Americans both directly (from "hunting" parties) and indirectly (from the railroads cutting across buffalo migratory routes). Destruction of the buffalo was encouraged by the government, as such action would take away the livelihood of the Indians. (4)(5)

Table B14.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Explain the events leading up to the Battle of Wounded Knee.

a) The Indian tribes of North America were caught up in a religious revival believing that an Indian Messiah was to appear and free them from the white intruders and the dead would rise. The white man called this the "Ghost Dance." The whites also saw this revival as a movement toward war again. Amid pressure from many people the Army tried to put a stop to this movement and a calvaly unit responded at Wounded Knee. (4)(5)

b) In 1889 the Indian tribes of the Northern Plains celebrated a "revival" called a "Ghost Dance." The dance wasn't a call to war, only a religious revival based on the belief that an Indian messiah was about to appear. With it came widespread popularity from other despairing tribes in the area: the dance quickly gained many followers. White settlers and miners panicked at the idea of an outbreak of war, and demanded the army be sent to quell the "activities." The 7th Cavalry, in accordance, came on the seen and quickly went to work. In 1890 in South Dakota, they arrested a band of Sioux men, women, and children and then the "battle" began. A dispute, then total insanity broke out. (4)(5)

c) In response to the "Ghost Dance" cult, a religious revival movement, which featured the coming of an Indian messiah who should return the buffalo and expel the white man, white miners and settlers, alarmed at what they feared might be another outbreak of warfare, demanded that the army put an end to this activity. (3)(5)

d) The Ghost Dance was a significant event that led upto the Battle of Wounded Knee. The Ghost Dance was part of a religious revival spurred by a prophet that had a vision of an Indian messiah about to appear to free the Indian people. Miners and settlers got wind of this and became afraid of a new Indian resistance movement. They demanded that the army keep a watchful eye on the Indians for signs of any trouble. Tensions were high and as a band of Sioux approached the Pine Ridge Reservation for food and shelter, they were shot down. (290 people). (5)(3)

g) The Battle of Wounded Knee was precipitated by a religious revival by the Sioux called the Ghost Dance. This revival was thought to be a call to war by the white settlers and eventually caused the settlers to demand that the cavalry step in and put a stop to it. (3)(5)

h) In South Dakota, as in other portions ofthe Plains, the Indians had been relocated to reservations by the late 1880s, and white settlers and miners had moved into much of the surrounding land. In 1889, a religious revival, the Ghost Dance, among the Indiands brought many whites to fear that an Indian uporising was imminent. The whites therefore demanded that the government send athe army to the Plains area to stop the dangerous religious activity. The army cam prepared to control the Indians both on and off the reservation. The everns can be explained by each group's unease at its position relative to the other group. The Indians wished to have their spaciuos ttraditional lands back free of white interference -- the central idea in their religius revival-- while the whites didn't feel free from the danger from the Indiands who had recently occuppied the area and had fought savagely to keep it. (5)(5)

Table B15.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: What was the most important theme or idea that was emphasized in this chapter?

a) The theme is that the white man broke promises made to the Indians in their (the white man's) quest to settle the plains. (5)(5)
b) That the Indians were constantly undermined by the repetitive advancements by the white man. The Indian losing his position and status. (4)(4)
c) That the American Indian War constantly victimized the Indians in this period by expansionistic Federal-white man goals. (4)(5)
d) The most important theme is that while the settling of the frontier by the whites was prosperous and helped to build a great nation, it extracted a heavy toll on the Native American Indians. (4)(5)
e) The opening of the Great Plains to the American white settlers, at the expense of the Indian civilization. (5)(5)
f) That the Indians were continually made promises that were then broken. (4)(3)
g) The most important theme seemed to be the white man's unfair and often brutal treatment of the Indians which occurred while settling the Plains. (4)(3)
h) The most important theme was how, and why, the Indaians were making their last stand on the Great Plains. (4)(4)

Table B16.

A sample of the responses of social studies methods students to and the scores of the two peer graders on an item on the 22-item take-home examination.

ITEM: Why is your answer to number 19 above the most important theme or idea in this chapter?

a) It is the most important theme because it describes the meanings of why the white man and the Indian reacted the way they did to the events. (5)(5)
b) Because it occurs or shows up in almost every section and paragraph in the chapter. It is an on-going theme, with events to back it up continuously. (5)(5)
c) Because it must be acknowledged that the plight of the American Indians at the hands of the Federal Govt did happen. (4)(5)
d) It shows that while it helped build our country it also helped to destroy a race of people and its culture and traditions. (4)(5)
e) The focus of this chapter was not only to tell the "History" of this period, but also describe factors which contributed to the theme (e.g., the buffalo, the railroad, the settlers). These events and factors converged to tell the "Story" which emerged as the theme. (3)(5)
f) Because it is mentioned the most often and it underlies the description of various events. (5)(5)
g) Throughout every subunit or idea expressed in the chapter, the undercurrents emphasize this treatment. (4)(4)
h) All of the other ideas tie into this theme at some point and most of the narraTive deals directly with that main theme. (4)(4)

APPENDIX C

TABLES PROVIDING RESULTS OF STATISTICAL ANALYSES NOT INCLUDED IN THE MAIN TEXT

Table C1.
Mean, standard deviation, and minimum and maximum scores by the two peer graders for individual items and total test for the 11 preservice social studies teachers in Group A.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.18	1.10	2.00	5.00	0.23
Q2	3.77	0.92	1.00	5.00	0.20
Q3	4.32	0.78	3.00	5.00	0.17
Q4	4.18	0.91	2.00	5.00	0.19
Q5	4.55	0.67	3.00	5.00	0.14
Q6	4.50	0.51	4.00	5.00	0.11
Q7	4.23	0.61	3.00	5.00	0.13
Q8	3.86	0.83	3.00	5.00	0.18
Q9	4.14	0.99	1.00	5.00	0.21
Q10	4.45	0.67	3.00	5.00	0.14
Q11	3.64	1.18	2.00	5.00	0.25
Q12	4.00	1.23	1.00	5.00	0.26
Q13	4.00	1.51	0.00	5.00	0.32
Q14	3.82	1.44	0.00	5.00	0.31
Q15	4.14	1.04	1.00	5.00	0.22
Q16	4.64	0.58	3.00	5.00	0.12
Q17	4.05	0.79	3.00	5.00	0.17
Q18	4.27	0.83	3.00	5.00	0.18
Q19	4.09	0.68	3.00	5.00	0.15
Q20	4.32	0.72	3.00	5.00	0.15
Q21	4.05	0.79	3.00	5.00	0.17
Q22	3.73	1.28	0.00	5.00	0.27
SUM20	83.14	9.22	62.00	97.00	1.97
SUM22	90.91	9.29	70.00	106.00	1.98
GRADEGPA	32.73	6.27	20.00	43.00	1.34

Table C2.

Mean, standard deviation, and minimum and maximum scores by the two peer graders for individual items and total test for the 25 preservice social studies teachers in Group B.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.76	0.66	2.00	5.00	0.09
Q2	4.36	1.19	0.00	5.00	0.17
Q3	4.30	1.30	0.00	5.00	0.18
Q4	4.50	1.09	0.00	5.00	0.15
Q5	4.32	0.91	2.00	5.00	0.13
Q6	4.44	0.88	2.00	5.00	0.13
Q7	4.40	1.05	0.00	5.00	0.15
Q8	4.10	1.27	0.00	5.00	0.18
Q9	4.40	1.05	0.00	5.00	0.15
Q10	4.04	1.51	0.00	5.00	0.21
Q11	4.34	1.21	0.00	5.00	0.17
Q12	4.74	0.80	0.00	5.00	0.11
Q13	4.38	1.40	0.00	5.00	0.20
Q14	4.24	1.48	0.00	5.00	0.21
Q15	4.74	0.63	3.00	5.00	0.09
Q16	4.70	0.74	2.00	5.00	0.10
Q17	3.74	1.78	0.00	5.00	0.25
Q18	4.66	0.94	1.00	5.00	0.13
Q19	4.70	0.74	2.00	5.00	0.10
Q20	4.38	1.12	1.00	5.00	0.16
Q21	4.54	1.16	0.00	5.00	0.16
Q22	2.52	2.36	0.00	5.00	0.33
SUM20	88.24	10.88	60.00	100.00	1.54
SUM22	95.30	12.43	65.00	110.00	1.76

Table C3.

Mean, standard deviation, and minimum and maximum scores by the two peer graders for individual items and total test for the 18 preservice social studies teachers in Group C.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.28	0.94	2.00	5.00	0.16
Q2	4.17	0.85	2.00	5.00	0.14
Q3	4.31	0.95	1.00	5.00	0.16
Q4	4.47	1.06	0.00	5.00	0.18
Q5	4.56	0.81	2.00	5.00	0.13
Q6	4.53	0.74	3.00	5.00	0.12
Q7	4.39	0.80	3.00	5.00	0.13
Q8	4.44	0.88	2.00	5.00	0.15
Q9	4.25	0.87	2.00	5.00	0.15
Q10	4.31	0.89	2.00	5.00	0.15
Q11	3.78	1.24	0.00	5.00	0.21
Q12	4.50	0.74	3.00	5.00	0.12
Q13	4.36	0.99	0.00	5.00	0.17
Q14	3.44	2.06	0.00	5.00	0.34
Q15	4.58	0.94	1.00	5.00	0.16
Q16	4.75	0.60	3.00	5.00	0.10
Q17	3.39	1.79	0.00	5.00	0.30
Q18	4.47	1.23	0.00	5.00	0.21
Q19	3.89	1.09	0.00	5.00	0.18
Q20	4.00	1.33	0.00	5.00	0.22
Q21	3.92	1.57	0.00	5.00	0.26
Q22	3.31	1.95	0.00	5.00	0.33
SUM20	84.86	8.60	67.00	99.00	1.43
SUM22	92.08	10.24	67.00	109.00	1.71

Table C4.

Mean, standard deviation, and minimum and maximum scores by the two peer graders for individual items and total test for the 22 preservice social studies teachers in Group D.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.66	0.75	2.00	5.00	0.11
Q2	4.34	1.24	0.00	5.00	0.19
Q3	4.50	1.02	0.00	5.00	0.15
Q4	4.48	1.13	0.00	5.00	0.17
Q5	4.34	0.89	1.00	5.00	0.13
Q6	4.05	1.31	0.00	5.00	0.20
Q7	4.25	1.06	1.00	5.00	0.16
Q8	4.25	1.04	0.00	5.00	0.16
Q9	4.32	1.12	1.00	9.00	0.17
Q10	4.25	1.14	0.00	5.00	0.17
Q11	3.61	1.81	0.00	5.00	0.27
Q12	4.84	0.48	3.00	5.00	0.07
Q13	4.34	1.52	0.00	5.00	0.23
Q14	4.00	1.64	0.00	5.00	0.25
Q15	4.68	0.80	1.00	5.00	0.12
Q16	4.68	0.91	1.00	5.00	0.14
Q17	3.75	1.88	0.00	5.00	0.28
Q18	4.25	1.46	0.00	5.00	0.22
Q19	4.39	0.81	2.00	5.00	0.12
Q20	4.50	1.00	1.00	5.00	0.15
Q21	4.20	1.72	0.00	5.00	0.26
Q22	3.11	2.17	0.00	5.00	0.33
SUM20	86.48	15.29	22.00	100.00	2.31
SUM22	93.80	17.63	22.00	110.00	2.66

Table C5.

Mean, standard deviation, and minimum and maximum scores by the two peer graders for individual items and total test for the 8 preservice social studies teachers in Group E.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	5.00	0.00	5.00	5.00	0.00
Q2	4.94	0.25	4.00	5.00	0.06
Q3	4.94	0.25	4.00	5.00	0.06
Q4	4.88	0.50	3.00	5.00	0.12
Q5	4.69	0.70	3.00	5.00	0.18
Q6	4.69	0.48	4.00	5.00	0.12
Q7	4.50	1.03	1.00	5.00	0.26
Q8	4.63	0.89	2.00	5.00	0.22
Q9	4.00	1.86	0.00	5.00	0.47
Q10	4.50	0.97	2.00	5.00	0.24
Q11	4.38	1.41	0.00	5.00	0.35
Q12	4.88	0.50	3.00	5.00	0.12
Q13	4.94	0.25	4.00	5.00	0.06
Q14	4.50	1.10	2.00	5.00	0.27
Q15	4.88	0.34	4.00	5.00	0.09
Q16	4.94	0.25	4.00	5.00	0.06
Q17	4.06	1.34	1.00	5.00	0.34
Q18	4.50	0.97	2.00	5.00	0.24
Q19	4.69	0.60	3.00	5.00	0.15
Q20	4.69	0.60	3.00	5.00	0.15
Q21	4.75	0.45	4.00	5.00	0.11
Q22	3.50	2.10	0.00	5.00	0.52
SUM20	93.19	6.73	76.00	100.00	1.68
SUM22	101.44	7.92	82.00	110.00	1.98

Table C6.

Mean, standard deviation, and maximum scores by the two peer graders for the individual items and total test for the 84 preservice social studies teachers in the five methods classes combined.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.58	0.82	2.00	5.00	0.06
Q2	4.29	1.07	0.00	5.00	0.08
Q3	4.42	1.03	0.00	5.00	0.08
Q4	4.48	1.03	0.00	5.00	0.08
Q5	4.44	0.84	1.00	5.00	0.06
Q6	4.39	0.94	0.00	5.00	0.07
Q7	4.35	0.95	0.00	5.00	0.07
Q8	4.23	1.06	0.00	5.00	0.08
Q9	4.27	1.12	0.00	9.00	0.09
Q10	4.25	1.16	0.00	5.00	0.09
Q11	3.94	1.43	0.00	5.00	0.11
Q12	4.63	0.81	0.00	5.00	0.06
Q13	4.37	1.31	0.00	5.00	0.10
Q14	3.98	1.64	0.00	5.00	0.13
Q15	4.63	0.81	1.00	5.00	0.06
Q16	4.72	0.71	1.00	5.00	0.05
Q17	3.74	1.67	0.00	5.00	0.13
Q18	4.45	1.15	0.00	5.00	0.09
Q19	4.36	0.88	0.00	5.00	0.07
Q20	4.35	1.07	0.00	5.00	0.08
Q21	4.27	1.36	0.00	5.00	0.11
Q22	3.10	2.10	0.00	5.00	0.16
SUM20	86.86	11.50	22.00	100.00	0.89
SUM22	94.23	13.10	22.00	110.00	1.01

Table C7.

Mean, standard deviation, and minimum and maximum scores by the two peer graders for the individual items and total test for the 29 secondary social studies students.

VARIABLE	MEAN SCORE	STANDARD DEVIATION	MINIMUM SCORE	MAXIMUM SCORE	STD ERROR OF MEAN
Q1	4.71	0.70	2.00	5.00	0.09
Q2	4.29	1.12	0.00	5.00	0.15
Q3	4.29	1.24	0.00	5.00	0.16
Q4	4.45	1.08	0.00	5.00	0.14
Q5	4.34	0.91	2.00	5.00	0.12
Q6	4.47	0.78	2.00	5.00	0.10
Q7	4.38	0.99	0.00	5.00	0.13
Q8	4.03	1.21	0.00	5.00	0.16
Q9	4.34	1.02	0.00	5.00	0.13
Q10	4.12	1.39	0.00	5.00	0.18
Q11	4.21	1.22	0.00	5.00	0.16
Q12	4.64	0.87	0.00	5.00	0.11
Q13	4.40	1.34	0.00	5.00	0.18
Q14	4.14	1.46	0.00	5.00	0.19
Q15	4.67	0.69	3.00	5.00	0.09
Q16	4.67	0.73	2.00	5.00	0.10
Q17	3.78	1.68	0.00	5.00	0.22
Q18	4.55	0.96	1.00	5.00	0.13
Q19	4.52	0.80	2.00	5.00	0.11
Q20	4.29	1.08	1.00	5.00	0.14
Q21	4.55	0.94	0.00	5.00	0.12
Q22	2.83	2.23	0.00	5.00	0.29
SUM20	87.29	11.12	60.00	100.00	1.46
SUM22	94.67	12.27	65.00	110.00	1.61

Table C8.

Means, standard deviations, minimum and maximum scores for differences in scores on the same exam by two peer secondary school social studies students who graded the same examination.*

VARIABLE	MEAN DIFFERENCE	STANDARD DEVIATION	MINIMUM DIFFERENCE	MAXIMUM DIFFERENCE	STD ERROR OF MEAN
DQ1	0.59	0.91	0.00	3.00	0.17
DQ2	0.66	0.77	0.00	2.00	0.14
DQ3	0.66	0.81	0.00	3.00	0.15
DQ4	0.62	1.24	0.00	5.00	0.23
DQ5	0.90	0.82	0.00	3.00	0.15
DQ6	0.79	0.77	0.00	2.00	0.14
DQ7	0.62	0.78	0.00	3.00	0.14
DQ8	0.97	0.91	0.00	3.00	0.17
DQ9	0.83	1.20	0.00	5.00	0.22
DQ10	0.79	0.86	0.00	3.00	0.16
DQ11	0.97	1.27	0.00	5.00	0.24
DQ12	0.52	0.95	0.00	4.00	0.18
DQ13	0.38	0.62	0.00	2.00	0.12
DQ14	0.76	0.99	0.00	4.00	0.18
DQ15	0.38	0.73	0.00	2.00	0.14
DQ16	0.66	0.94	0.00	3.00	0.17
DQ17	1.00	1.22	0.00	5.00	0.23
DQ18	0.55	0.91	0.00	4.00	0.17
DQ19	0.48	0.91	0.00	3.00	0.17
DQ20	0.66	1.04	0.00	4.00	0.19
DQ21	0.34	0.55	0.00	2.00	0.10
DQ22	0.90	1.37	0.00	5.00	0.25
DSUM20	9.97	8.61	0.00	34.00	1.60
DSUM22	10.45	8.47	0.00	34.00	1.57
SMABDF20	13.76	8.75	2.00	38.00	1.62
SMABDF22	15.00	9.19	2.00	40.00	1.71

*N = 29 students.

APPENDIX D

SAMPLES OF DEFINTIONS PROVIDED BY SOCIAL STUDIES PRESERVICE TEACHERS TO TERMS FOUND IN TYPICAL TEST ITEMS AND TASK ASSIGNMENTS PROVIDED PRE-COLLEGE LEVEL SOCIAL STUDIES STUDENTS

Table D1.

Copy of the form used to collect definitions from the participants.

Defining Terms

Listed below are a number of terms that are commonly found on assignments, tests, and materials used in the classroom. After each word, write a definition of what that term means you need to write down to get a correct answer. For example, an item on a test may ask you "to describe (or explain) the Boston Tea Party." In this situation, what would it mean for you"to describe?" or "to explain?"

Please take your time and define each term as completely as possible. What you write down may help us to design more worthwhile assignments and better test questions.

1. **Identify**

2. **Describe**

3. **Explain**

4. **Understand**

5. **Comprehend**

6. **Compare**

7. **Contrast**

8. **Define**

9. Analyze

10. Discuss

11. Cause

Table D2.

A sample of the definitions the pre- and in-service teachers provided for each of the eleven terms they were asked to define following their grading of the take home exam.

Identify
-- pinpoint person/place/thing and state what part they played in what is being studied
-- noting specific features
-- to pick out parts of a story or characters
-- to name or state who someone is or what something is
-- to catagorize
-- to categorize things
-- to know and explain
-- to show that you know what it is

Describe
-- list the attributes of
-- express something or something
-- explaining the setting and structure and plot of what happened as completely as possible
-- to give details which tells about someone, something or an event
-- explaining the details of items or ideas
-- give in detail
-- to explain in terms of appearances, characteristics of, etc.
-- to list and explain specific details about a specific item
-- show understanding by explaining what it is

Explain
-- using the attributes to show comprehension
-- discuss and write about a situation
-- discuss whatever happened as completely as possible
-- to give details, "reasons" to describe a concept
-- discussing the meaning or reason
-- discuss the meaning of something
-- to describe
-- to state in one's own words details about an event, object, or act.
-- discuss or write the meaning or reason for something

Understand
-- mostly deals with facts
-- comprehending and knowledge of material
-- to be able to talk and write about the story or whatever easily
-- to find meaning in a concept. You learn what something means.
-- comprehend
-- too similar to comprehend. Just have to use one or the other.
-- to comprehend
-- to have in one's own mind a good picture of a particular concept.
-- to have a good idea about what they are doing

Comprehend
-- understand the reasons behind the facts
-- recalling the understanding of material
-- understand material that was read
-- you are able to go a step further from understanding. You can apply a concept, you can compare and contrast, find relationships, etc.
-- understand
-- too similar to understand. Just have to use one or the other.
-- to understand
-- to understand completely
-- to understand the meaning of

Compare
-- mostly listing opposite and the same attributes
-- take two situations and what is different or the same
-- write differences between two different things explaining each
-- to find similarities between two concepts, find their relationships
-- the similarities of 2 or more items, or concepts
-- similarities
-- to list the similarities
-- to list likenesses
-- describe the similarities of two or more things

Contrast
-- view of the opposite points of a specific topic
-- same as above for compare
-- describe the difference to the item or the opposite of it
-- to find differences between two concepts and find their relationships
-- the differences between 2 or more items or concepts
-- differences
-- to list the differences
-- to list differences
-- describe the differences between two or more items

Define
-- how is a specific word or phrase used in the context of what has been studied
-- educated meaning
-- a statement describing something
-- to give the definition or tell what a word or concept means
-- the literary meaning of a word or concept
-- brief explanation
-- to give meaning to
-- state characteristics and specifics
-- to give the meaning of a word or phrase

Analyze
-- delve into the reasons - interpretations of facts
-- discuss a situation
-- organized thoughts of occurrances and your opinion of them
-- to take a "in depth look" about a concept, find fault or good, you often incorporate your beliefs and opinions into analysis.
-- making an observation and trying to come to a conclusion
-- tell what you think
-- to break down into sections and discuss
-- to compare, contrast, and summarize
-- think over an object, idea, and come to a hypothesis

Discuss
-- converse or write in narrative form what was learned about the topic
-- talk or write about a situation
-- answer the question as best you know and add or comment to the statement at hand
-- to answer questions, find relationships, cause-effect, consequences, sequence, etc. Cover all aspects of a concept so you fully understand it.
-- talk about certain concepts, values, ideas, beliefs, etc.
-- to talk about in detail
-- talk about in general terms
-- talk about an object, idea -- give many different views

Cause
-- the reasons/feelings/happenings behind an action or happening
-- explain why something happened
-- why an occurrence happened
-- to state why something happened has occurred, events that led up to cause "it" to happen
-- the reason for an event happening
-- events leading to cause something
-- the purpose
-- reasons for the way something is